FIGURE 101

GGGCCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTCGCAGANCACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGA

702 1237

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCCTGTCCCCGAA

104/237

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCCTGCGAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCTG

FIGURE 106

TTCCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCTATGACCCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTCTTACCCACG

107/237

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTCNGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGAA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCCTCCAGCGCGCCCACCCGGTAGAGGACCCCCGCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAAACTTAAAGCGGGCGCAGCATTAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAAC**ATG**GCGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGCTCCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAATAAAATTATATCCACAACTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTG**TAG**AGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATA
TTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTCAGTAAATAACAACTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

109/237

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTDRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMLQPFDYDPNEKSKHKFMVQSMFAPTDTSDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:
amino acids 224-239

N-glycosylation site.
amino acids 68-71

N-myristoylation site.
amino acids 59-64, 64-69 and 235-240

1101 237

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTTAACTTTAAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

112 1237

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCCGGCCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTG
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANTCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTCATCTCATGTTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAA**ATG**GAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTCAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGCTGCCCACTGTGTTCATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAAATCTGGCCGGGGTC
AGAGGATTGCCGAAGGGAGGCCTTCCTTTCAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACTTGGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTCAGAAGGAC
TACAACGTTGCTGTTCGCATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTCACGG
GAACGATGCCAATTGTGCTTACGGC**TAA**CAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTTGCCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

718/237

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:
amino acids 1-16

N-glycosylation sites.
amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.
amino acids 236-239

Serine proteases, trypsin family, histidine active site.
amino acids 165-170

119/237

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGC**ATG**GAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCG**TAA**GTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAAAAAA

720/237

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQELEEGAAVSEEPS

Important features:

Signal peptide:
amino acids 1-25

N-glycosylation site.
amino acids 251-254

Thrombospondin 1
amino acids 385-399

von Willebrand factor type C domain proteins
amino acids 385-399, 445-459 and 42-56

721237

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AG**ATG**GCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCGCCGGTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAACTGGCTAACATTCTTTT
TACCAGGGAACTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTCATGGGCTTTTTTCAAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTCAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAA**TAG**GAACAAGGAGTAAAAGAGCTGTTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTCAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
<subunit 1 of 1, 336 aa, 1 stop
<MW: 36865, pI: 9.15, NX(S/T): 2
MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELRQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTVNVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:
amino acids 1-21

Short-chain alcohol dehydrogenase family protein
amino acids 134-144, 44-56 and 239-248

N-glycosylation site.
amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTTGGGATAAGAGAATTTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

724/237

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACC**ATG**CATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTT**TAA**AAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:
amino acids 1-29

Endoplasmic reticulum targeting sequence.
amino acids 403-406

Tyrosine kinase phosphorylation site.
amino acids 203-211

Thioredoxin family proteins
amino acids 50-66

726237

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCACGCGTCCG**ATG**GCGTTCACGTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCT**TAG**AACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAACTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACTTGTA
GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCCTTTAGT
GCAATACAATAAAACTCTGAAATTAAGACTC

729/239

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, pI: 5.60, NX(S/T): 0
MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:
amino acids 1-20

Type II transmembrane domain:
amino acids 11-31

Other transmembrane domain:
amino acids 57-77 and 123-143

7301237

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAG**ATG**GCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAAT**TTA**AGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pI: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:
amino acids 1-31

Transmembrane domain:
amino acids 241-260

N-glycosylation site.
amino acids 90-93

733/237

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGC
TGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCA**ATG**GGCCGCGGCTG
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAAACTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCCTTTG
CATTTTGATGAGAATTCATTTTTTGCTGGGGATAAAAAAGAAGCACACAAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACT
TCAGGAACTTGTTACAGAATATTCAT**TAA**AGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCATTCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

736/237

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLKRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:
amino acids 1-23

N-glycosylation site.
amino acids 280-283 and 384-387

Amidation site.
amino acids 94-97

Glycosaminoglycan attachment site.
amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate
amino acids 216-222

Interleukin-7 proteins
amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGC**ATG**GAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC
ATGGTGAGC**TAA**GGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAAAAAAAAATACAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTC
ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMVS

Important features:

Signal peptide:
amino acids 1-22

Cell attachment sequence.
amino acids 70-73

N-glycosylation site.
amino acids 98-101

Integrins alpha chain proteins
amino acids 67-81

740/237

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCTTCCAACTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

141/237

FIGURE 141

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

742 1237

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:
amino acids 1-29

Transmembrane domain:
amino acids 230-255

N-glycosylation site.
amino acids 40-43 and 134-137

Tissue factor proteins.
amino acids 92-119

Integrins alpha chain proteins
amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTTGGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGG**ATG**GGGGCTGCAGCTCGTCTGAGCGCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAACTGCGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTCAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGC**TGA**GACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 177-199

N-glycosylation site.
amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins
amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC
GTTCGTGTCCCCGCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCCACCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGCGTATTCCTGCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
CACCTCTCCCAGGAAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCCTGGGGAA
ATCCTGAGGTCATTCATT**ATG**AAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGCTGTTCATGGCTAGA
GCAATTCCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAAACTTTTGGAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTCGTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCATAACATGAACATCTGGGGGCAG
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAAACTGGTGGGGCCATGCCCCTTACAAA
CATGGGCGGCCCTGTTCTGCTTGCCCACCTAGTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAATGTCCTGCTGGCTGTTTGGAT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTCATCAAGTCCAATAGA
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTG
ACTTGTGAAACAACTGTGGAACAGCTCTGTCCATTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCCT
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAATTGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
AAGACCTACATTGCTTCTTTTCAGAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATTC
AGAGTGTTTGCTGTTGTG**TGA**AACTGAATACTTGGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
ATTTTGTATAAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAAAGGTGCCAAATGCATA
TAAATCTTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTCATATATTATATGGTGCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCC
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
TGTTAATTTAGGCATATAGAATATTAAATTCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC
CAAATGAATCTGTTAAAATGTTTGATTCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
GAACAAAGATGAACTAATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
AATCTGTAAAATGTTAGTTTTGGTAATTTTTTTTCTGCTGGTGGATTTACATATTAAATTTTTTCTGCTGGTGGA
TAAACATTAAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLEKYMDEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC
HNMNIWGQIWPKAVYLVCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKAFRV
FAVV

Important features:

Signal peptide:
amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein
amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site
amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCCGCTCCGACGGGCCAGCGCCCTCCCC**ATG**TCCCTGCTCCCACGCCG
CGCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAA**TAG**GGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA
GGAACAATGAGCTTGGTGGACACATTTCATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCCT
GGTTAATTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNAWNEKRRVYEE

Signal sequence:

amino acids 1-34

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTCACCCCGGTCTGCGTC**ATG**TTAAACTCCAATGTCCTCCTGTG
GTTAACTGCTCTTGCCATCAAGTTCACCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTA
TGCCTCACCCCCCACTGGAGAGAGGCGGTTTCAGCCCCCAGAACCCCCGTCCTCCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGT
GCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAA
CATGATTGACGGCAGCATTTTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
AGGGTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTCAGCCTGTTGACCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGC
CCTGTCCAGCTGGGCAGTGAACTACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCATCCCAGACGACCCCCAGATCCTGATGGA
GCAAGGCGAGTTCCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCTGAAGTTCGTGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCCAACGACTTTGACTTCTCCGTGTCCAACTTCGTGGACAACCTTTACGG
CTACCCTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAACCC
GGAGACGCGGCGGAAAACCCTGGTGGCTCTCTTTACTGACCACCAGTGGGTGGCCCCCGCCGTGGCCGCCGACCT
GCACGCGCAGTACGGCTCCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCCAGCTG
GGCAGATTCGGCCCATGGTGATGAGGTCCCCTATGTCTTCGGCATCCCCATGATCGGTCCCACCGAGCTCTTCAG
TTGTAACTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTCATGACCTACTGGACGAACTTCGCCAAAACTGG
TGATCCAAATCAACCAGTTCCTCAGGATACCAAGTTCATTCACACAAAACCCAACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGGTTGGAACTCGTTCCTCATTTGCACAACTTGAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCCTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
AACCACCAAACGCCCAGCAATCACTCCTGCCAACAATCCCAAACACTCTAAGGACCCTCACAAAACAGGGCCTGA
GGACACAACTGTCCTCATTGAAACCAAACGAGATTATTCCACCGAATTAAGTGTCACCATTGCCGTCGGGGCGTC
GCTCCTCTTCCTCAACATCTTAGCTTTTGCGGCGCTGTACTACAAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA
CTACACCCTCACGCTGCGCCGGTCGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA
CACACTGACGGGGATGCAGCCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTCCACCACTAGAGTA**TAG**CTTTGCCCTATTTCCCTTCCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTTGTCCCACT
GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCCTTATCGTTGGTGTTTTCCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTTAAAATAACTGCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAAGGCCCCGGGTGTTTCCAACGTCATGGAAGCAGCT
GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGCCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAAAGAATGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAAGAG
GTGTTTTGCCAGCCTGAACTATATTTAAGAGACTTTGT

751/237

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVPYV
FGIPMIGPTELFSCNFSKNDVMLSAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTIAVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:
amino acids 1-24

Transmembrane domains:
amino acids 189-204, 675-692

152/237

FIGURE 152

GGGAAAG**ATG**GCGGCGACTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTCGAGTACTTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCCTGGTGATTCGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAACTGACAGTGGAGAGAACCCCAGAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTTTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTAC**TGA**
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCGTTGGGTT
TTGCATTTGACCCAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACT
TCTTCCCTGCCTTACCTTCCTTTCACTCCATTCATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAAGCAACATTTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACTCTGAGTCCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

753/237

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:
amino acids 1-38

Transmembrane domain:
amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATA**ATG**CTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTCATTCGCCACGGAGACAGGTACCCACTGTATGTCATTCCCA
AAACAAAGCGACCAGAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACTGGAAGCTTTCA
TTAGTCACATGTCAAAAGGATCCGGAGCCTCTTTCGAAAGCCCCTTGAACTCCTTGCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAAACTCCTGCCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCACTGGGAAAA
GCCGGACCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCGCTGTTCTGCTCTGGAAGCTGCTATTGCCCGGTAAGAAACCAGTATCTGGAAAAGGAGC
AGCGTCGTCAGTACCTCCTACGTTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGTACTTCGGGTATTCTCTCCTGGGTGCCCACCCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTGGGCCTTTCAGAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCCAGTGAACATTCCGTCCGGATTCTTTACAATGGCGTCGATGTCACATTCCACACCT
CTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTCACAGGGAAGGATTC**TAA**AAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATTGCTTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTTAATCTTAGACATT
TTTACCTTGTCCTTGTTAAGAATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTCAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTCATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCCTGGGCTGTAGTTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCCTCCAGTAACTTCTGCTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAATATT
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTAAATGCGTTTTTGGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACTAATGGTGCTAATTCAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTCATCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTTGCATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAGAATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF

Signal sequence:
amino acids 1-18

156/237

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
GCCGCTGTTCACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGA**ATG**AT
TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACGCTATTCAGTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCT
GAATCCGCGCGCAGCGGCAGCTTGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTTCGTGAAAGTGAATTAGAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT
CCCTCTACCCCACGCCTGGGATCCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGAACACTCA
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCCGAATTGGTGCTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACCTGGTCCTTACGGCCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGCG
CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCAGCGTTTGCTCAGCCCGAGTACCGCGCGAGCGTTCC
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCCAAGTTTTCAAACTAGATTGTAATTCAGGGACAATATC
AACAATAGGGGAGTTGGACCACGAGGAGTCAGGATTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATA
TTCTGCGCGAGCCAAAGTCCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCCTCACCTCTCT
CGCCAGCTCGGTTCCCGAAAACTCTCCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTTCATCCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA
TAGTTTAGTCACAGACATAGTCTTGGATAGGGAACAGGTTCCTAGCTACAACATCACAGTGACCGCCACTGACCG
GGGAACCCCGCCCCTATCCACGGAAACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCCGGTCTT
CCCTCAGGCCTCCTATTCCGCTTATATCCCAGAGAACAATCCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
CGACCCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
CTTGCAAGTGAAAGTGATGGCGCGGGACAACGGGCACCCGCCCCTCAGCAGCAACGTGTCGTTGAGCCTGTTCGT
GCTGGACCAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCCTCCCCACGGACGGTTCCACTGGCGTGGAGCT
GGCTCCCCGCTCCGCAGAGCCCGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
CTGGCTGTCCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCACGGCCAGCC
CCCTCTCTCCGCCACTGTCACGCTCACCGTGGCCGTGGCCGACAGCATCCCCCAAGTCCTGGCGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG
CGTCTTCCTGGCCTTCGTCATCTTGCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTGCTGCAGGC
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTCGCACTTTGTGGGCGTGGACGGGGTGCAGGCTTTCCTGCAGAC
CTATTCCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACCTGATCTTCCCCCAGCCCAACTATGCAGA
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAAGCGAGCCCCTTTTGCTGTCAGGTGATTCGGTATTTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTTTTTTTTAATTGCTCTGTCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAACTCCTAGGCTCAAGCAATTATCCCACCTTTGCCT
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCACGCCTGTAATCCCAGTACTTTGGGAGGC
CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAGACCAGCC**TGA**CCAACATGGAGAAACCCCGTCTATACTAA
AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT
TGCTTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG
AAACTCTATCTCA

157/232

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILLLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAFLQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGSLTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:
amino acids 1-30

Transmembrane domains:
amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACC**ATG**GGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGC**TGA**TTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

759/237

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 51-71

N-glycosylation site.
amino acids 110-113

Serine proteases, trypsin family, histidine active site.
amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.
amino acids 182-188

Kringle domain proteins motif
amino acids 205-217

760/237

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCGC
CGCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCGC
GCCCGCGCCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCCGCCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCCTCGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCGCCCAACCCCTACG**ATG**AAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCCTGCACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCCACGGCACATCAATGACTCACCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTCACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGC**TGA**CCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCCTGCCGCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTCCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

761237

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGSGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:
amino acids 1-26

Leucine zipper pattern.
amino acids 135-156

Glycosaminoglycan attachment site.
amino acids 436-439

N-glycosylation site.
amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain
amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTAC
AGTCCCACAGAACCGTCCTCCCAGGAAGCTGAATCCAGCAAGAACA**ATG**GAGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTTTCCTTTCTCCTTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTTGTCAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
TTTCCAGAGGGAACAAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTCACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCCTCCTGGG
ACTACGTTTCCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTCGGGTCCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCCTATTCACTTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACTAAAAAAACAACTCGATTTCGAAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTCAGTGTTTCAGATCTTGATTCAGGAG
AAAATGGGAAAATTAGTTGCTCCATTCAGGAGGATCTACCCTTCCTCCTGAAATCCGCGGAA
AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTCGTC
CGCGAGAACAACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTCACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTCGGCGTGTGGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCTTTTCGGTGCTCCTGTTC
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGTCGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCCTATCCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCAG**TGA**CCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:
amino acids 1-26

Transmembrane domain:
amino acids 685-712

Cadherins extracellular repeated domain signature.
amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).
amino acids 285-292

N-glycosylation site.
amino acids 418-421, 436-439, 567-570 and 786-789

169237

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCA**ATG**GTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTCGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAG**TAG**TGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHMVMLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:
amino acids 1-25

Transmembrane domain:
amino acids 321-340

Homologous region to dilsufide isomerase
amino acids 212-302

N-glycosylation site.
amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281 and 293-296

Thioredoxin domain
amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCCACG**ATG**CTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGC**TA**
GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTCACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACTAATA
AAATCATGAATATTTTAA

167237

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:
amino acids 1-20

Cysteine rich domain, homolgous to frizzled N terminus
amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACG**ATG**GCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTCGAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCCACCCCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCGA
GTGTTCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCCTGGGGCCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCCTCGCTAGAGGACCCCAGAGGGCCCGGGAGCCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACG**TAG**CTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

1691237

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGILIMLAMTALTVRRSEPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:
amino acids 1-44

Transmembrane domains:
amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268, 280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.
amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.
amino acids 468-471

FIGURE 170

GTCCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCCATCATTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCCTTGGGGTCCTGCTTGTTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAGC**ATG**ATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATTGTCAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTCAGAGGAAGTGCC
ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAACTGGGCCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTCAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGGA
GCTGGAAATCTCTGAGAGCGCCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCC
TAACACCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTTGCCTTGGATGTCATTGTGGGCCCTGATGAGAC
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAAATCCATTCATTTTTTGATCTGGTGTTAACTGC
CTATGACAATGGGAACCCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG
CCCTGCGTTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAAACT
GACCGCCACAGACCCTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCCTCAGTAAGCACATGCCTCCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCATTCTGCGTCGACCTCTAGACTATGAAAAGAACCC
TGCCTACGAGGTGGATGTTCAGGCAAGGGACCTGGGTCCCAATCCTATCCCAGCCCATTGCAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTCACATGGGCCTCCCAGCCATCACTGGTGTCAGAAGC
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCATGGCAGATGACTTGGATTCAGGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGCCCAAATATACCCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACTTACCCTCTCTTCACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTCAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATT
CCTTTTGACAACCATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCTCTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTCGTCAATGTCACCAATGCCAGCAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCCCTTACAGACCCGAGCCCTGTT
GAGGGTCATGTTTGTCACCAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTCGAT
GCTGACGGTGATCTGCCTGGCTGTACTGTTGGGCATCTTCGGGTTGATCCTGGCTTTGTTCATGTCCATCTGCCG
GACAGAAAAGAAGGACAACAGGGCCTACAACTGTCGGGAGGCCGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
CCAGAAACACATTCAGAAGGCAGACATCCACCTCGTGCCTGTGCTCAGGGGTCAGGCAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT
CCACCTCACCCCGACCCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCCAGGCAGAGGAATGCCTCCCGGGAGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCCTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCCTGAGAAAGAATCAGGGCCCCGTCAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTCGCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTCAGCAAATCTCCCAGCTGCT
GTCCTTGCTGCATCAGGGCCAATTCCAGCCCAAACCAAACCACCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCCTCAC
CACCAACTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCGG
CAAGGCAGAGGCACCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCCTCAGTTTAGACTTGGCCACCAGTGCAGCCTCAGGCATGAAAGTGCAAGGGGACCCAGGTGG
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGTGCCTG**TGA**ACATACCTCAGACGCCT
CTGGATCCAAGAACCAGGGGCCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAG
CGGCGGCCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCCTTGTAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGTT
TCCTGGCAAAACATATGTGGAGCACAAAGGGTCAGTCCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAA

771/237

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 amd 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

172237

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCC**ATG**AGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAAACTAGGAGATCCCTTTCCCATTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT
TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAAACTAA
AGGTGAAAAATACACTGGAACTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTAAGGTTCACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPFPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, 425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

174/237

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTCTC

FIGURE 176

CTCGCGCAGGGATCGTCCC**ATG**GCCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTCGAACTGCTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCCTGGGGGCAAGATTGTTACCTGTGCACACCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCTTCTCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCCCGCGCCAACCA
CAAGGGTGCTGTGGTCATCCTGCGCAAGGACAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTCGGGATCAGCCTGGCTGTCCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTTGTCGCCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTCAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCCTGAGCCGTAACCTGGAAGAACCCAAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCCTCAATGCCCACCAGCCCAGCACCCAGCGGGCAGAGATCCACTTCCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTCAGCGACACGGAATTCCAACC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCCTGGAGCTGAT
GGTCACCAACCTGCCATCGGACCCAGCCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCCAGCTCCTGGT
CATGCTTCCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCCAGGTCACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAACTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCAGCAACTCTTCTTCTCTGGTGTGGTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTCACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCTGG
GCAGAAAGGGCTTTGCTCTCCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCCAGCATGTCCTGGTGGCCAGTGTCCTCTGC
TGAGAAGAAGAAAAACATCACCCTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGCTGCCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCCTCCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTGATCCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGCTAGCACTGCTGGTGCTGCTCCTGTGGAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACTGGGGCAGCCCCCGGCGGGAGGGCCCGGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCC**TAG**GTTCC
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTCGCATCAAGATTTGGCAGGATCGGCTTCCTCAGGGGCACAGACCTCTCCCACCCACAAGAAC
TCCTCCCACCCAACTTCCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGTCCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCCTCTCCCATTCACCCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAAAAAATATTTGAGAACAAAAAAAAAAAAAAAAAAA

771237

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGSAFLNIM
WPHEIANGKWLLYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMAVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

771/237

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAACAAG**ATG**CTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCGGACGGCGGTAATTTTC
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTGCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC
CATGATGTATACATT**TGA**TTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFDQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPI
CKDSLGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:
amino acids 1-16

Leucine zipper pattern.
amino acids 246-267

N-myristoylation sites.
amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins
amino acids 353-365 and 339-352

780/237

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTCAGCCCGGGATCGCCCCAGCAGGG**ATG**GGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCTT
CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTCAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACT**TAA**ACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCATTAATGGTCTTCTCCAAAATATTTTGAGATATA
AAAGTAGGAAACAGGTATAATTTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
TTTCCTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAAATGCGGCAGTTACAAATTAACTGTGGAAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
AGATTTTTATAACCAAATACATTTCAGTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
CCAAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAAGCTTCCTCCTCCATAGGGACACATTTTCTCTAACCCTTAACTAAAGTGTAGGA
TTTTAAAATTAAATGTGAGGTAAAATAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTTAGAAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AAACTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTCTGGTTTGAATTATTTTTGTAAGCAGGTACATTTTATA
AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAAAAATTTTTATTATAACAACTTTTAT
TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAATTTAACTTCATCATGAATGAACTATAATTTTTAAGTTATGCCCATTTATAACGTTGTTTAT
GACTACATTGTGAGTTAGAAACAAACTTAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
TCTTAGGCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
ATGATCAATTACCTTAATTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCCTATGAATACCCCTCCCTTTTGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACTTCTTACCATATAAAAACGATAATTGCTT
TATTTGGAAAAGAATTTAGGAATACTAAGGACAATTATTTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
CATAACCAAAAAAGCAAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC
CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTCATGTGAATGTCAAGACATGGAGGGTGTTTGTAATTTTATGGTAAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAAATTGACAAAAAATGAGCACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTTAT
GTGAAATTTTAAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
CTCAAACTGCTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
AAATTATCAAAGGAAAA

TM 1237

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 195-217

N-myristoylation site.
amino acids 43-48

Tyrosine kinase phosphorylation site.
amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACT**ATG**CTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGAC**TAG**GGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCCTTTTTCTTCACCTTCATTTCAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:
amino acids 1-26

C-type lectin domain signature.
amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACC**ATG**CGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTG**TAG**AGTCCCCGCCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

785237

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:
amino acids 1-17

N-glycosylation site.
amino acids 46-49

7861737

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAG**ATG**ATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTCGCGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGA
GACAC**TAA**ACCAGCTATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCCTCTTGTTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAAAAAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pI: 8.85, NX(S/T): 1
MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:
amino acids 1-23

N-glycosylation site.
amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain
amino acids 110-126

788/237

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

789237

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGG
CTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTCAC
TTAAATCAGAACTTGCATAAGAAAGAGA**ATG**GGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTGG
AGTGTCCAAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTCACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGCTGCTGGTATTGGCTGGCTGATCACTTTTTGTTCAAA
AGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAACTACTTTCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACT
AAAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGACTGTTCCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATCATGGAAA
GAAGATTCTATATGATATACTTGCCTTTGCCAAAGAAAGTGTGAATTCTCATGTTACCACGCTTGGACCTCAAAA
TTTTCCTGCCAATGACAAAGAACCATGGCTTGTTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTCATGA
GGGACTCTGTAACATGTATAACATTCAGGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAACTAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCCTTGCCAAGTCTTAATGCCAGAATGGAAAAGAATGGCCCGGACATTAACTGGACTGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATTCCCTGAGAAT
CTGGGGTCTAGGATTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAACTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAAAAATTGGAAACTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTT**TGA**TAATGTTGAAGATGAAGAAAAAGTTTAAAAGAAATTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGGTTCACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAACAACCTTTAAAAAATATTAAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAACT
TAGTTTTTGGTCACTTGTTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTCGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCCTTCTCA
AAGGTTGAAAAAATGCTTTTAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATTCTGTAGTTTGCTGTATCATCCAGGAAAACCTGAGGGAAAAAAATTA
TAGCAATTAACTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTCATGTATTTTCTGAAATTGCTTTCATAGAAATTTTCCCACTGATAGTTGATTTTTGAGGCATCTAATAT
TTACATATTTGCCTTCTGAACTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTCATAGTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTCCAGTTACTAGTTTACTGTCAGAGGGCTGCCTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGA
CTCAAAGAATCACAAATTTGTCAGTAACATGTAGTTGTTTAGTTATAATTCAGAGTGTACAGAATGGTAAAAATT
CCAATCAGTCAAAAGAGGTCAATGAATTAAAAGGCTTGCAACTTTTTCAAAAAAAAAAAAAAAAA

7901237

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIHHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.
amino acids 744-747

Cytochrome c family heme-binding site signature.
amino acids 158-163

Nt-dnaJ domain signature.
amino acids 77-96

N-glycosylation site.
amino acids 484-487

791237

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCC**ATG**AACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCACCA
AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAA**TGA**ATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTCACATTTTT
TCAGTCCTGATAATATTAAAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGTCTGTTTTTCCTT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTTCCTCCCTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

192/237

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:
amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein
amino acids 165-202, 37-49, 112-122 and 210-219

793/237

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGC
AGG**ATG**ACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGCACACGGGGCCGCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGCACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGAC**TGA**AAGCTTTCGCGTTGCTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, pI: 9.54, NX(S/T): 4
MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELTA
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAAQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:
amino acids 1-31

N-glycosylation sites.
amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein
amino acids 329-332

195/237

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGG**ATG**GCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGGGCTGCGGCAGTGTTCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCC**TGA**TTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTCACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.
amino acids 138-141

Short-chain alcohol dehydrogenase family protein
amino acids 10-22, 81-91, 134-171 and 176-185

797/237

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGA**ATG**GACTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCG
CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTC**TGA**ATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

198237

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:
amino acids 1-20

N-glycosylation site.
amino acids 75-78

Homologous region to IL-17
amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTCACAGTGACTCTTGGAATT
CCTGCTTGGTGTTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TG**TAG**ATTTAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTCCTGAATGAGCAAGCTTCTCTTAAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTCAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTTCCATCTTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:
amino acids 1-22

Transmembrane domain:
amino acids 191-211

N-glycosylation site.
amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)
amino acids 56-72

Flavodoxin proteins
amino acids 173-187

207/237

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAGGGGCTCTCAAGAACA**ATG**GAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTCACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATG**TAA**GAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTCGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTAGT
AGAGACAGGGTTTCTCCATGTCGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTAAATACA
AAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAAACTATATTTCCTCATG
TGCTGATTTGTCTTATTTTTTTCATACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTCAT
TTCAATAAAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

202/237

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:
amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 197-200

N-myristoylation sites.
amino acids 35-40 and 151-156

Homologous region to LDL receptor
amino acids 34-67 and 70-200.

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTCGGAGGAAA**ATG**ACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCCTGCTGAGTCTGCTCTTCCTGGTCCAAGGTGCCCACGGCAGGGGCCACAGG
GAAGACTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCTTTCCCTGCAGCCCACCCTGCTTCCCGA
TCCTTCCCTGACCCCAGGGGCCTCTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCCGCTGTTAGCCACTTCTGTCACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCCAGTGCC
GCCAGCTTCACCTTCTCCTTCCACAGTCCTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCATGGGGGACATGGTGTCCTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCCACA
TTGAGCAGCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCACTTGACCTACTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTCGTCTCTGCCCTGGCCTGCCTTGTCACCATTGCCGCCTACCTCTGCTCCAGG
GTGCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC
CTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCCTGCACTTCTCCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGC
CTGGTGTTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGCTGGGCCTCAGCCTGGTCCTTGGCCTGCCCTGGGCCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTCGTCCTCTACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATC**TAG**GCCTCCAGCCCACCTGCCCATGTGATGAAG
CAGAGATGCGGCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCAGCCGCAGACT
TTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGG
GGCCAGGCCTTGGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACCCTGAGGGCACTCTGCATCCTCTGTCATTTTAACCTCAGGTGGCACCCAGGG
CGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCCAGGCCCCCTCCCTCCCTCAGCCCCCCAGTCCTCCCTCCATCTTCCCTGGGGTTC
TCCTCCTCTCCCAGGGCCTCCTTGCTCCTTCGTTCACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA
GTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG
CTCACCCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCCTAAGAGGTTCTCTCCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCCATCATGGT
TAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACTCCTCCTGCCAACATTCAGTCTGGTATGTGAGGCGTGCGTGAAGCAAGAACTCCTGGAGCTACAGGGACA
GGGAGCCATCATTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAGTACATGTTCATTGTAGAGA
ATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590 and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327 and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTTTCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAACTACCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCCAGTCAGCCAA**ATG**AGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCATGCCACAATTCACAGAAGCTGTGAGTCCTTCAGTGCTACCCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGA**TAG**CCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAATATTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

208/237

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGAAGA
AGACTAAAA**ATG**GTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTATCCTTTTTAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAA
GATGCCTTCCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCCATTTCCTTGTGCGCCG
TGTAAAAATAATTCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAAAGTTTTACGTCTA
CACAGTAACTCTCTTCAGCATGTGCCCCCAAGATGGTTTAAGAACATCAACAAACTCCAGGAACTGGATCTGTCC
CAAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCCAGGTCCTGGAACAATTACATTATTTCAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCTCTGATTTTCAGCATCTTTCTTTCCTCAAATGCCTG
AATCTGTCAGGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
CTGCAGAAACTGATGATGAACGACAATGACATCTCTTCCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT
ATGCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAAACTCCAGTGT
CTAAAGAACCTGGAAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACTGTTCC
AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA
CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTCCTTGAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTC**TAG**CCCTTCTTTGCAAAACACAACTGCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYR
NPCYVSYSIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:
amino acids 1-26

Transmembrane domain:
amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAACATGGAAAACATGTTCCTTC
AGTCGTCAATGCTGACCTGCATTTTCCTGCTAATATCTGGTTCCTGTGAGTTATGCGCCGAAGAAAATTTTTCTA
GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTCATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC
CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCCTCAACCTAAAAAACCTAAGGGAGTTACTGC
TTGAAGACAACCAGTTACCCCAAATACCCTCTGGTTTGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAACA
ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAACTGCT
ATTTTAACAAAGTTTGCGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
CACTATCTTTCAATTCTCTTTCACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACA
CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAACT
GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTGCTTCAATTAATATAGATCGTTTTG
CTTTTCAAAACTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAACGATGCTGCCCCGCTTAGAAATACTTGACTTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCATTTAAGAGGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCCTGATGCAGCTTCCAAACTTATCGACTATCAACTTGGGTATTAATTTTATTA
AGCAAATCGATTTCAAACTTTTCCAAAATTTCTCCAATCTGGAAATTATTTACTTGTCAGAAAACAGAATATCAC
CGTTGGTAAAAGATACCCGGCAGAGTTATGCAAATAGTTCCTCTTTTCAACGTCATATCCGGAAACGACGCTCAA
CAGATTTTGAGTTTGACCCACATTCGAACTTTTATCATTTCACCCGTCCTTTAATAAAGCCACAATGTGCTGCTT
ATGGAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTCATTGGGCCAAACCAATTTGAAAATCTTCCTGACA
TTGCCTGTTTAAATCTGTCTGCAAATAGCAATGCTCAAGTGTTAAGTGGAACTGAATTTTCAGCCATTCCTCATG
TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATTGTCCGACTTGG
AAGTTCTAGATCTCAGCTATAATTCACACTATTTCAGAATAGCAGGCGTAACACATCATCTAGAATTTATTCAAA
ATTTCACAAATCTAAAAGTTTTAAACTTGAGCCACAACAACATTTATACTTTAACAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTAGAATTAGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAATCTGACACGTCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATG
AAGCATTCCTTAATTTGCCAGCGAGTCTCACTGAACTACATATAAATGATAATATGTTAAAGTTTTTTAACTGGA
CATTACTCCAGCAGTTTCCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAAACTACTCTTTTTAACTGATAGCC
TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCCACCTACCCTCTGGCTTTC
TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
AAACTAAGACCACCACCAAATTATCTATGTTGGAACTACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAAATTCCCAGACTGGTAGATGTCATTTGTGCCAGTCCTG
GGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTCAGATGTCACTGCAGTGATATTAT
TTTTCTTCACGTTCTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTTGTTTTACTGGGATGTTT
GGTTTATATATAATGTGTGTTTAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACTTTCTATGATG
CTTACATTTCTTATGACACCAAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG
AGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
TCATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACTTTA
AAACAGCTTTTTACTTGGCTTTGCAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCCTGCTGGAGC
CAGTGTTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA
ACCCGAAGGCAGAAGGCTTGTTTTGGCAAACTCTGAGAAATGTGGTCTTGACTGAAAATGATTCACGGTATAACA
ATATGTATGTCGATTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTCGCGCCATAATAAAGATGCAAAG
GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTATCCCAAAACTTAGTGGTTTAAAACAACACA
TTTGCTGGCCCACAGTTTTTGAGGGTCAGGAGTCCAGGCCCAGCATAACTGGGTCCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTCACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCCACAAGGCAGCTTGCTTC
ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGTAATCGAATCAAAAAAGTGAT
ATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACCACAGGTCCCACCAGCTCCATGGGAGTGACCACC
TCAGTCCAGGGAAAACAGCTGAAGACCAAGATGGTGAGCTCTGATTGCTTCAGTTGGTCATCAACTATTTTCCCT
TGACTGCTGTCCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGCCACTGTGACTAATGGTCCTAATA
TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTCGGTTTTAT
TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTTTAAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA
TTGTTAATTGCCATTGCTGTAAATCTTAAAATGAATGAATAAAAATGTTTCATTTTACAAAAAAAAAAAAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDLEFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFSNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDLS
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:
amino acids 1-26

Transmembrane domain:
amino acids 826-848

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTG
GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGA
CTCG**TGA**CTGCCCAGCGCTCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCC**ATG**AGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCT
GGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCCTGCCGTCAGCTGCCGGCCG
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCGCGTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCCTTTTCCTCATCCTGGGGACATGTACAC
TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTCAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCCTCGGGCGCTACCAGATG
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
TCAAGAATTTCCAGATAAACAACCAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGGCCTCCCATTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTG
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAACTG
TTCTAGAAGTCCTCATTTGCTTCTTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTCATACTTTCCTCGTGG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCCTTGCCCCCCAGTGTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCCTTGGACCCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTCAC
TGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCACAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTTTTTCCTTTTGACTCCTGCTCCCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
ACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAGCCT
CGGCTGAGTGAGGGAAAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCCTAAGGCCTCGGCCCACCTCTGGCT
ATGGTAACCACACTGGGGGCTTCCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTT
CACCCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTCAGGGAAGAAGATTTATGT
ATTATATGTGGCTATATTTCCTAGAGCACCTGTGTTTTCCTCTTTCTAAGCCAGGGTCCTGTCTGGATGACTTAT
GCGGTGGGGGAGTGTAAACCGGAACTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLICFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGA
AAAAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCGACTCCACC**ATG**AACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCTTACA
GAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGGCTGGATTCGGAGGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACTGGGGATGCTGCTGGGTGGGCGGCCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCACCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACCAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATC
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACTCCCGTGACTTCCTGCGG
CACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGG**TAG**ACCTGGATCAGGGGA
GAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG
TGCCTCTGCTTTGGGGGTGCCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGQLCEVW

Type II Transmembrane domain:
amino acids 32-57

FIGURE 226

GCCCGGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCGCTCCCGCGCCCTCCTCCCTCCCTCCTCCC
CAGCTGTCCCGTTCGCGTCATGCCGAGCCTCCCGGCCCCGCCGGCCCCGCTGCTGCTCCTCGGGCTGCTGCTGCT
CGGCTCCCGGCCGGCCCGCGGCGCCGGCCCAGAGCCCCCCGTGCTGCCCATCCGTTCTGAGAAGGAGCCGCTGCC
CGTTCGGGGAGCGGCAGGTAGGTGGGCGCCCGGGGGAGGCGCGGGCGGGGAGTCGGGCTCGGGGCGAGTCAGCGC
CAGCCCGGAGGGGGCGCGGGGCGCAGGTGGCTCGGCGCGGCGGGCGGCCCGGAGGGTGGGCGGGGGCAGAAGGGC
GCGGTGCCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGCTGGGCAGCGGCCTCCAGC
CAAGCCCGTCCCCGCAGGCTGCACCTTCGGCGGGAAGGTCTATGCCTTGGACGAGACGTGGCACCCGGACCTAGG
GGAGCCATTCGGGGTGATGCGCTGCGTGCTGTGCGCCTGCGAGGCGCAGTGGGGTCGCCGTACCAGGGGCCCTGG
CAGGGTCAGCTGCAAGAACATCAAACCAGAGTGCCCAACCCCGGCCTGTGGGCAGCCGCGCCAGCTGCCGGGACA
CTGCTGCCAGACCTGCCCCCAGGACTTCGTGGCGCTGCTGACAGGGCCGAGGTCGCAGGCGGTGGCACGAGCCCG
AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTT
CTCAGACTCCAATGGCAGTGTCCTGTTTGAGCACCCTGCAGCCCCCACCCAAGATGGCCTGGTCTGTGGGGTGTG
GCGGGCAGTGCCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCA
CCCTTCAGGGGAGGTCTGGGGGCCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGAC
TCTAGAAGGCCCCCACCAGCAGGGCGTAGGGGGCATCACCCTGCTCACTCTCAGTGACACAGAGGACTCCTTGCA
TTTTTTGCTGCTCTTCCGAGGCCTTGCAGGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACCAGGGGCA
GCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCCAGGAACCAGGCTTTGCTGAGGTGCTGCCCAACCTGACAGT
CCAGGAGATGGACTGGCTGGTGCTGGGGGAGCTGCAGATGGCCCTGGAGTGGGCAGGCAGGCCAGGGCTGCGCAT
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGTCCTGCAAAGTGTCCTTTGTGGGGCTAATGCCCTGATCCC
AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCCTGATCCTCCAGGTGCAATT
GGTAGGGACAACCAGTGAGGTGGTGGCCATGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCCCACTGTCCT
GTGCCACATGGCTGGCCTATCCTCCCCTGCCCCCAGGCCGTGGGTATCTGCCCTGGGCTGGGGTGCCCGAGGGGC
TCATATGCTGCTGCAGAATGAGCTCTTCCTGAACGTGGGCACCAAGGACTTCCCAGACGGAGAGCTTCGGGGGCA
ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGCCCCTAGCAGGAGCCCTGGTGCTACC
CCCTGTGAAGAGCCAAGCAGCAGGGCACGCCTGGCTTTCCTTGGATACCCACTGTCACCTGCACTATGAAGTGCT
GCTGGCTGGGCTTGGTGGCTCAGAACAAGGCACTGTCACTGCCCACCTCCTTGGGCCTCCTGGAACGCCAGGGCC
TCGGCGGCTGCTGAAGGGATTCTATGGCTCAGAGGCCCAGGGTGTGGTGAAGGACCTGGAGCCGGAACTGCTGCG
GCACCTGGCAAAAGGCATGGCTTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCT
CTCCTCCCAGGTGCACATAGCCAACCAATGTGAGGTTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAGGGGGT
GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCGCCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCGC
CAAACCTGGTGGTCCTGGGCGGCCCCGAGACCCCAACACATGCTTCTTCGAGGGGCAGCAGCGCCCCCACGGGGC
TCGCTGGGCGCCCAACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGATCTGTGACCCGGT
GGTGTGCCCACCGCCCAGCTGCCCACACCCGGTGCAGGCTCCCGACCAGTGCTGCCCTGTTTGCCCTGGCTGCTA
TTTTGATGGTGACCGGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCCTTTGGCTTAATTAA
GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCAGTGTCCCCGGCTGGC
CTGTGCCCAGCCTGTGCGTGTCAACCCCACCGACTGCTGCAAACAGTGTCCAGGTGAGGCCCACCCCCAGCTGGG
GGACCCCATGCAGGCTGATGGGCCCCGGGGCTGCCGTTTTGCTGGGCAGTGGTTCCCAGAGAGTCAGAGCTGGCA
CCCCTCAGTGCCCCCGTTTGGAGAGATGAGCTGTATCACCTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGT
GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCA
CTGTGTCCCAGTGCCTCTGGGGGACACTCAGTGTCTGCTCTGTCTTGTACCAGGCAGGGGTGCCTCACTGTGAGC
GGGATGACTGTTCACTGCCACTGTCCTGTGGCTCGGGGAAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCCACC
GGCGGCGTAAGTGAGGGAGTCCAGGGTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGTGGGACTCCTGATCAG
GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAATCCACCCTCACAGCA
ACCTGGTGGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTTCTGAAGCTCAGAGAAATTAAGCAACGAG
ATGAAGGTCACCCAGCTGTGTGCACTGACCTGTTTAGAAAATACTGGCCTTTCTGGGACCAAGGCAGGGATGCTT
TGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCACTCCCTCTCCCCTCCTCCAACATTCCCTCCCTTCTGTCTCC
AGCAGCCCCAGAGACCAGAACTGATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTTAGGGAGCAGCCAGAGGGCC
AAGTGACCAAGAGGATGGGGCCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCATTCTCCTGTGGGAAG
CCCAGTGCCTTTGCTCCTCTGTCCTGCCTCTACTCCCACCCCCACTACCTCTGGGAACCACAGCTCCACAAGGGG
GAGAGGCAGCTGGGCCAGACCGAGGTCACAGCCACTCCAAGTCCTGCCCTGCCACCCTCGGCCTCTGTCCTGGAA
GCCCCACCCCTTTCTTCCTGTACATAATGTCACTGGCTTGTTGGGATTTTTAATTTATCTTCACTCAGCACCAAG
GGCCCCGGACACTCCACTCCTGCTGCCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTTGTATTTATTAAAAC
ATTTCTTTTTCAGTCTTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCCTGGTGGAGAAGGG
GCNGAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
CGTGGCNNTTGGCTGGCATNCCTGGGTTCCGCAGAGGGGCTGGGGATGGTTCTTGAGATGGTCTAGAGACTCAAG
AATTTAGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCCACATCGCTGGCCTGTTTGCTGACTTCATG
TTTGAAGTTGCTCCAGAGAGAGAATCAAAGGTGTCACCAGCCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTTCT
TTCCCTCCCCTCCCCTCCCCTCCCCTCCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTCGGA
GGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCT
CCTTTCTTTCTTCTTTTCTTCCTCCCTCCCTCCCTTTCTTCCCCTTTTCCTTCCTTCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTCG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTCGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

228/237

FIGURE 228

ATGCCTACTACCTTCCAACTAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCTCACGGGCTACCCCCACCTCACCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAAGAGAGAAACAACACCCAACAACTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAATCCTGTGGCGCGCCGCCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGAC**ATG**ATG
CTGTTGGTGCAGGGTGCTTGTTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCGGAAAATGCTGTGTCATT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTCAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTCACCAGCATATTCTACCTGAAGAATGCCATTCTACAA**TAA**A
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTCATGTTTGTAATCTACAACTTTTCAAAAGCATTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTCAGTTCCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAACTAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACCAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTC
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCCTTTCACTAGTCCAAGCCAAAAA
TTTTAAGATGATTTGTCAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCCATCTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGDTAVLRCYLED
GASKGAWLNRSSIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNVDVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFPDVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLSSFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:
amino acids 1-31

Transmembrane domain:
amino acids 326-345

N-glycosylation sites.
amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.
amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.
amino acids 178-186

N-myristoylation sites.
amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297, 302-308, 319-325

Myelin P0 protein:
amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCA**ATG**AGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCCTGGAGATGTATTTCCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACTTCACTCAGAACCCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCTGGCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCAAAAGGGGAAACTGGAACTAAGGGAGAGAAAGGAGACCTGGGTCTCCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCATGGGGCCTCCTGGAGCCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTTTTCCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGCAGTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACTCAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCAGTGTCGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTC**TGA**CCCGGAAACCCTTTCA
CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDLGLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAKGDQGQPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGQQGRKGESGVPGPA
GVKGEQGSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352, 400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCAAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAA**ATG**AAACTGATGGTACTTGTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCG
TGATTCCTTGCAACAATCAA**TGA**GAATCTTCATGTATTCTGGAGAACACCATTCCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:
amino acids 1-20

N-glycosylation site.
amino acids 72-76

Tyrosine kinase phosphorylation site.
amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAG**ATG**
TGGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAACAAAACTTTCAGCTTGCAAAGCAAATTCAATCCCAGTGGAAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGTCCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
GTATTCCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAAACTTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
CTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA
GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
ACTAAAAAAAAGTCCTTCCCCAGAGTTCAGTGGCATGCCCAGGATAAGCAAATTGGGATCTGG
AAATGATTTTGAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTCG
AGGAGGGATGGTGTTTGAGCTAGCCAATTCCATAGTGCTCCCTTTTGATTGTCGAGATTATG
CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGAC
AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCC**TAA**GAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIARYGKVFRGNKVKNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMG
GSAPPDSSWRGSLKVPYNVGPGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDSLFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:
amino acids 1-40

N-glycosylation sites.
amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340, 459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.
amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.
amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341, 360-366, 427-433, 529-535, 707-713

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 September 2000 (14.09.2000)

PCT

(10) International Publication Number
WO 00/53756 A3

(51) International Patent Classification[7]: C12N 15/12, 15/63, 5/10, 1/21, 1/19, C07K 14/47, 14/705, 19/00, 16/18, 16/28, G01N 33/68, 33/543, A61K 47/48, 51/08, 38/17

(21) International Application Number: PCT/US00/04341

(22) International Filing Date: 18 February 2000 (18.02.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PCT/US99/05028	8 March 1999 (08.03.1999)	US
60/123,957	12 March 1999 (12.03.1999)	US
60/126,773	29 March 1999 (29.03.1999)	US
60/130,232	21 April 1999 (21.04.1999)	US
60/131,445	28 April 1999 (28.04.1999)	US
60/134,287	14 May 1999 (14.05.1999)	US
60/141,037	23 June 1999 (23.06.1999)	US
60/145,698	26 July 1999 (26.07.1999)	US
60/162,506	29 October 1999 (29.10.1999)	US
PCT/US99/28313	30 November 1999 (30.11.1999)	US
PCT/US99/28551	2 December 1999 (02.12.1999)	US
PCT/US99/28565	2 December 1999 (02.12.1999)	US
PCT/US99/30095	16 December 1999 (16.12.1999)	US
PCT/US99/31243	30 December 1999 (30.12.1999)	US
PCT/US99/31274	30 December 1999 (30.12.1999)	US
PCT/US00/00219	5 January 2000 (05.01.2000)	US
PCT/US00/00277	6 January 2000 (06.01.2000)	US
PCT/US00/00376	6 January 2000 (06.01.2000)	US

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ASHKENAZI, Avi, J. [US/US]; 1456 Tarrytown Street, San Mateo, CA 94402 (US). BAKER, Kevin, P. [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). BOTSTEIN, David [US/US]; 2539 Somerset Drive, Belmont, CA 94002 (US). DESNOYERS, Luc [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). EATON, Dan, L. [US/US]; 75 Knight Drive, San Rafael, CA 94901 (US). FERRARA, Napoleone [US/US]; 2090 Pacific Avenue #704, San Francisco, CA 94109 (US). FILVAROFF, Ellen [US/US]; 538 18th Avenue, San Francisco, CA 94121 (US). FONG, Sherman [US/US]; 19 Basinside Way, Alameda, CA 94502 (US). GAO, Wei-Qiang [CN/US]; 641 Pilgrim Drive, Foster City, CA 94404 (US). GERBER, Hanspeter [CH/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). GERRITSEN, Mary, E. [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). GRIMALDI, Christopher, J. [US/US]; 1434 36th Avenue, San Francisco, CA 94122 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth, J. [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). KLJAVIN, Ivar, J. [US/US]; 3673 Crescent Drive, Lafayette, CA 94549 (US). KUO, Sophia, S. [US/US]; Apartment 3, 59 Surrey Street, San Francisco, CA 94131 (US). NAPIER, Mary, A. [US/US]; 1015 Hayne Road, Hillsborough, CA 94010 (US). PAN, James [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). PAONI, Nicholas, F. [US/US]; 1756 Terrace Drive, Belmont, CA 94002 (US). ROY, Margaret, Ann [US/US]; 2960 Webster Street #4, San Francisco, CA 94123 (US). SHELTON, David, L. [US/US]; 5845 Clover Drive, Oakland, CA 94618 (US). STEWART, Timothy, A. [US/US]; 465 Douglass Street, San Francisco, CA 94114 (US). TUMAS, Daniel [US/US]; 3 Rae Avenue, Orinda, CA 94563 (US). WILLIAMS, P., Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).

(74) Agents: KRESNAK, Mark, T. et al.; Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 00/53756 A3

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

(88) Date of publication of the international search report:
1 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/04341

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N15/63 C12N5/10 C12N1/21 C12N1/19
C07K14/47 C07K14/705 C07K19/00 C07K16/18 C07K16/28
G01N33/68 G01N33/543 A61K47/48 A61K51/08 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 57983 A (ZYMOGENETICS INC) 23 December 1998 (1998-12-23) the whole document ---	1-3, 5-13, 15-21
A	KLEIN R D ET AL: "Selection for genes encoding secreted proteins and receptors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, no. 93, 1 July 1996 (1996-07-01), pages 7108-7113, XP002077277 ISSN: 0027-8424 the whole document --- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 July 2000	18. 10. 00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/04341

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YOKOYAMA-KOBAYASHI M ET AL: "A signal sequence detection system using secreted protease activity as an indicator" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 163, no. 2, 3 October 1995 (1995-10-03), pages 193-196, XP004041983 ISSN: 0378-1119 the whole document ---	
A	EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document ---	
A	WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document ---	
P,X, L	WO 99 46281 A (BAKER KEVIN P ;CHEN JIAN (US); GENENTECH INC (US); GURNEY AUSTIN ()) 16 September 1999 (1999-09-16) whole document, particularly the pasasages referring to PRO213 (e.g. pages 2,50,123,183), and the claims ---	1-3, 5-13, 15-21
P,X	WO 99 54437 A (MILLENNIUM BIOTHERAPEUTICS INC) 28 October 1999 (1999-10-28) the whole document -----	1-3, 5-13, 15-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/04341

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Invention 1 : claims 1-3, 5-13, 15-21, all partially

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 1: claims 1-3,5-13,15-21, all partially

Nucleic acid with seq.ID.1, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.2 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.2 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide.

Invention 2: claims 1-29,38-43, and 50-53, all partially

Nucleic acid with seq.ID.611, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.612 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.612 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO337 (seq.ID.523) using its interaction with PRO4993 (seq.ID.612), method for linking a bioactive molecule to a cell expressing PRO337 through the use of PRO4993, and method of modulating at least one activity of said cell thereby.

Invention 3: claims 1-29,38-43, and 50-53, all partially

Nucleic acid with seq.ID.522, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.523 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.523 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO4993 (seq.ID.612) using its interaction with PRO337 (seq.ID.523), method for linking a bioactive molecule to a cell expressing PRO4993 through the use of PRO337, and method of modulating at least one activity of said cell thereby.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 4: claims 1-21,30-37,44-49, and 54-57, all partially

Nucleic acid with seq.ID.613, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.614 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.614 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO700 (seq.ID.523), PRO725 (seq.ID.616) and/or PRO739 (seq.ID.618) using their interaction with PRO1559 (seq.ID.614), method for linking a bioactive molecule to a cell expressing PRO700, PRO725 or PRO739 through the use of PRO1559, and method of modulating at least one activity of said cell thereby.

Invention 5: claims 1-21,30-37,44-49, and 54-57, all partially

Nucleic acid with seq.ID.89, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.90 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.90 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO1559 (seq.ID.614) using its interaction with PRO700 (seq.ID.90), method for linking a bioactive molecule to a cell expressing PRO1559 through the use of PRO700, and method of modulating at least one activity of said cell thereby.

Invention 6: claims 1-21,30-37,44-49, and 54-57, all partially

Nucleic acid with seq.ID.615, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.616 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.616 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO381, as represented by seq.ID's 144 and 145,
PRO386, as represented by seq.ID's 149 and 150,
PRO540, as represented by seq.ID's 156 and 157,
PRO615, as represented by seq.ID's 161 and 162,
PRO618, as represented by seq.ID's 168 and 169,
PRO719, as represented by seq.ID's 177 and 178,
PRO724, as represented by seq.ID's 182 and 183,
PRO772, as represented by seq.ID's 189 and 190,
PRO852, as represented by seq.ID's 195 and 196,
PRO853, as represented by seq.ID's 205 and 206,
PRO860, as represented by seq.ID's 210 and 211,
PRO846, as represented by seq.ID's 215 and 216,
PRO862, as represented by seq.ID's 220 and 221,
PRO864, as represented by seq.ID's 225 and 226,
PRO792, as represented by seq.ID's 230 and 231,
PRO866, as represented by seq.ID's 235 and 236,
PRO871, as represented by seq.ID's 244 and 245,
PRO873, as represented by seq.ID's 253 and 254,
PRO940, as represented by seq.ID's 258 and 259,
PRO941, as represented by seq.ID's 263 and 264,
PRO944, as represented by seq.ID's 269 and 270,
PRO983, as represented by seq.ID's 283 and 284,

Claim : 9

PRO1057, as represented by seq.ID's 295 and 296,
PRO1071, as represented by seq.ID's 300 and 301,
PRO1072, as represented by seq.ID's 302 and 303,
PRO1075, as represented by seq.ID's 308 and 309,
PRO181, as represented by seq.ID's 321 and 322,
PRO195, as represented by seq.ID's 329 and 330,
PRO865, as represented by seq.ID's 336 and 337,
PRO827, as represented by seq.ID's 345 and 346,
PRO1114, as represented by seq.ID's 351 and 352,
PRO237, as represented by seq.ID's 357 and 358,
PRO541, as represented by seq.ID's 362 and 363,
PRO273, as represented by seq.ID's 369 and 370,
PRO701, as represented by seq.ID's 374 and 375,
PRO704, as represented by seq.ID's 379 and 380,
PRO706, as represented by seq.ID's 384 and 385,
PRO707, as represented by seq.ID's 389 and 390,
PRO322, as represented by seq.ID's 394 and 395,
PRO526, as represented by seq.ID's 399 and 400,
PRO531, as represented by seq.ID's 404 and 405,
PRO534, as represented by seq.ID's 409 and 410,
PRO697, as represented by seq.ID's 414 and 415,
PRO717, as represented by seq.ID's 419 and 420,
PRO731, as represented by seq.ID's 424 and 425,
PRO218, as represented by seq.ID's 429 and 430,
PRO768, as represented by seq.ID's 436 and 437,
PRO771, as represented by seq.ID's 441 and 442,
PRO733, as represented by seq.ID's 446 and 447,
PRO162, as represented by seq.ID's 451 and 452,
PRO788, as represented by seq.ID's 453 and 454,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO1008, as represented by seq.ID's 455 and 456,
PRO1012, as represented by seq.ID's 458 and 459,
PRO1014, as represented by seq.ID's 463 and 464,
PRO1017, as represented by seq.ID's 465 and 466,
PRO474, as represented by seq.ID's 467 and 468,
PRO1031, as represented by seq.ID's 469 and 470,
PRO938, as represented by seq.ID's 471 and 472,
PRO1082, as represented by seq.ID's 476 and 477,
PRO1083, as represented by seq.ID's 482 and 483,
PRO200, as represented by seq.ID's 487 and 488,
PRO285, as represented by seq.ID's 495 and 496,
PRO286, as represented by seq.ID's 497 and 498
PRO213-1, as represented by seq.ID's 505 and 506,
PRO1330, as represented by seq.ID's 507 and 508,
PRO1449, as represented by seq.ID's 509 and 510,
PRO298, as represented by seq.ID's 514 and 515,
PRO403, as represented by seq.ID's 525 and 526,

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/04341

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9857983 A	23-12-1998	AU 7979898 A	04-01-1999
		EP 0996628 A	03-05-2000
EP 0834563 A	08-04-1998	JP 10179178 A	07-07-1998
		US 5824504 A	20-10-1998
WO 9707198 A	27-02-1997	US 5707829 A	13-01-1998
		AU 6712396 A	18-02-1997
		AU 6768596 A	12-03-1997
		CA 2227220 A	06-02-1997
		CA 2229208 A	27-02-1997
		EP 0839196 A	06-05-1998
		EP 0851875 A	08-07-1998
		JP 11510045 T	07-09-1999
		US 6043344 A	28-03-2000
		WO 9704097 A	06-02-1997
		US 6074849 A	13-06-2000
		US 5969093 A	19-10-1999
WO 9946281 A	16-09-1999	AU 3072199 A	27-09-1999
		AU 3075099 A	11-10-1999
		WO 9947677 A	23-09-1999
		AU 1532499 A	15-06-1999
		EP 1032667 A	06-09-2000
		WO 9927098 A	03-06-1999
		AU 3757099 A	08-11-1999
		WO 9954467 A	28-10-1999
		AU 1070399 A	10-05-1999
		EP 1025227 A	09-08-2000
		WO 9920756 A	29-04-1999
WO 9954437 A	28-10-1999	AU 3965799 A	08-11-1999